



Red Hat RFI Response- Topic 2

Solicitation NNH15ZDA012L: Preparation for the Development of a Community-based Roadmap for NASA's Planetary Data Services

Date of Submission: January 4, 2016

National Aeronautics and Space Administration Point of Contact:

Dr. Michael New

michael.h.new@nasa.gov

Red Hat Point of Contact:

Jon Shallant, Account Manager

Phone: 703.629.0564

jshallant@redhat.com

Treatment of Red Hat's Proposal

This response document contains Red Hat trade secrets, including but not limited to sensitive pricing information, and may not be disclosed to a third party without Red Hat's prior written consent. The exchange of information under this response shall be protected under any applicable agreement between the parties related to the treatment of confidential information. In the absence of such agreement, the recipient of this response shall protect the information contained herein in the same manner as it protects its own confidential information and may not be shared with any third party.

This is a proposal for a potential business relationship, and is neither a contract nor an offer that could be accepted by the other party. If the parties decide to enter into a business relationship, they will memorialize the terms of that relationship in a separate agreement between them or their respective business partners.

This document is Copyright 2015 by Red Hat, Inc. All rights reserved.

Contents

EXECUTIVE SUMMARY.....1

Red Hat Response.....2

SOLUTION OVERVIEW.....2

RED HAT GLUSTER STORAGE DIFFERENTIATORS.....2

 SOFTWARE-ONLY STORAGE.....2

 OPEN SOURCE SOFTWARE.....2

 SUPPORT FOR STORAGE CO-RESIDENT AND DATA LOCALITY AWARE APPLICATIONS.....3

 USER SPACE.....3

 MODULAR, STACKABLE ARCHITECTURE.....3

 DATA STORED IN NATIVE FORMATS.....3

 NO EXTERNAL METADATA SERVER.....3

 GLOBAL NAMESPACE TECHNOLOGY.....3

 STANDARDS-BASED FILE AND OBJECT STORE.....4

 CONCLUSION.....4

EXECUTIVE SUMMARY

Red Hat is the world's leading provider of open source solutions, using a community-powered approach to provide reliable and high-performing cloud, virtualization, storage, operating system, and middleware technologies. Open source is a model for software developed by a community of diverse developers, resulting in higher quality, more secure, more easily integrated software at a vastly accelerated pace.

With a Red Hat subscription, there are no hefty up-front licensing fees, no support incident limits, no upgrade costs, and no hidden charges. Instead, NASA will receive unlimited support, software binaries, product documentation, updates, upgrades, bug fixes, and security patches all at one low, predictable price.

Red Hat has a proven track record managing data and enabling end users to get faster access to data, as shown in the following three case studies:

1. Cern is the European Organization for Nuclear Research, a conglomeration that spans 34 countries, 140 facilities and 100K machines. They needed to send information across their network but their homegrown system was brittle, prone to losing messages and could not scale. They used Red Hat JBoss Fuse to allow for fast, resilient system for sharing data across the different machines to improve research and collaboration.

2. The Federal Aviation Administration (FAA) needed a reliable, cost effective means for data exchange between themselves, industry and airline partners. The platform needed to be easy for partners to work with, and thus had to accommodate numerous protocols, standards and data formats to allow for flexible yet secure data exchange. The platform must also simultaneously handle the increased capacity and future demand. The FAA utilizes Red Hat to provide all these capabilities to their strategic System Wide Information Management (SWIM) program while reducing interdependence on existing systems, making them more agile in their data exchange capabilities.

3. At the University of Reading, the Department of Meteorology needed a highly reliable, available, and scalable storage file system to manage data for its scientific research projects in weather, climate, and earth observation. With Red Hat Storage Server, the department now has an enterprise-grade product—backed by world-class service and support—that saves departmental IT staff valuable research time they used to spend on maintenance and administration tasks.

Red Hat and NASA have had an outstanding relationship for many years. Red Hat has presence at every NASA Center, and within many scientific and research organizations. NASA leverages not only Red Hat Enterprise Linux, but also many solutions from our Cloud, Middleware and Storage portfolio. Red Hat understands NASA's passion for open source solutions as we continue to support the organization in increasing productivity, enhancing security/compliance requirements and lowering the costs of doing business. Red Hat believes we are in the best position to meet many of the items listed in the RFI. We look forward to NASA's response and the possibility of adding value to this specific organization's mission.

Red Hat Response

SOLUTION OVERVIEW

During the last decade, enterprises have seen enormous gains in scalability, flexibility, and affordability as they have migrated from proprietary, monolithic server architectures to virtualized, open source, standardized, and commoditized servers.

Unfortunately, the same has not been true in the storage industry. Proprietary, monolithic, and scale-up solutions, which dominate the storage industry today, do not deliver the outcomes of a modern, software-defined datacenter. They do not deliver scalability, flexibility, or the economics that datacenters, workloads, and cloud computing environments need in today's hyper-growth, virtualized, and increasingly cloud-based world. Red Hat Gluster Storage addresses this gap.

Red Hat Gluster Storage is an open, software-defined storage solution for private, public, and hybrid cloud environments. Based on the open source project GlusterFS, Red Hat Gluster Storage provides a distributed scale-out file systems technology to meet the needs of unstructured, semi-structured and big data storage environments.

Red Hat Gluster Storage enables organizations to combine large numbers of commodity storage and compute resources into a high-performance and centrally managed pool of storage. Capacity and performance can scale linearly and independently on-demand, from a few terabytes to petabytes and beyond, using both on-premise commodity hardware and the public cloud compute and storage infrastructure. As such, Red Hat Gluster Storage will improve interaction between the PDS and data providers.

RED HAT GLUSTER STORAGE DIFFERENTIATORS

SOFTWARE-ONLY STORAGE

One of the cornerstone principles defining Red Hat Gluster Storage is that storage should be considered a software problem. Locking customers into one particular storage hardware vendor or one particular hardware configuration cannot solve today's storage problems. As a software-only storage solution, Red Hat Gluster Storage has been designed to work with a wide variety of industry-standard commodity storage, networking, and compute servers.

OPEN SOURCE SOFTWARE

Red Hat Gluster Storage delivers functionality by embracing the open source model. Based on the open source project GlusterFS, a distributed scale-out file system technology, Red Hat Gluster Storage takes many "upstream" projects in the community and packages it with Red Hat Enterprise Linux, for a stable and enterprise-grade storage offering. As a result, Red Hat Gluster Storage users benefit from a worldwide community of developers. These developers are constantly testing the product in a wide range of environments and workloads and providing continuous feedback.

SUPPORT FOR STORAGE CO-RESIDENT AND DATA LOCALITY AWARE APPLICATIONS

Red Hat Gluster Storage is ideal for running applications that need to run co-resident with storage, like Apache Hadoop map-reduce. These types of applications must avoid the extra network hop between the compute node and the storage server if possible to improve response times for analytics workloads.

The software-defined characteristic of Red Hat Gluster Storage makes it amenable to these workload types.

USER SPACE

Unlike traditional file systems, Red Hat Gluster Storage operates in user space. This significantly eases the process of installing and upgrading the product. It also enables expedited delivery of features that do not depend on kernel merge windows. And it means that users who contribute to GlusterFS do not need specialized kernel expertise.

MODULAR, STACKABLE ARCHITECTURE

Red Hat Gluster Storage is designed using a modular and stackable architecture approach. Configuring Red Hat Gluster Storage for highly specialized environments is a simple matter of including or excluding particular modules.

DATA STORED IN NATIVE FORMATS

With Red Hat Gluster Storage, data is stored on disks using native formats (e.g., XFS). The product has implemented various self-healing processes for data and therefore is extremely resilient. Furthermore, files are naturally readable without Red Hat Gluster Storage. Therefore, if a customer migrates away from Red Hat Gluster Storage, his/her data is still completely usable without any required modifications or data migration.

NO EXTERNAL METADATA SERVER

In a scale-out system, one of the biggest challenges is keeping track of the logical and physical data locations. Most distributed systems solve this problem by creating a separate index with file names and location metadata that usually resides in a server process, and is often referred to as a metadata server.

Unfortunately, this results in both a central point of failure and a huge performance bottleneck. Red Hat Gluster Storage does not create, store, or use a separate index of metadata that needs to be externally stored. Instead, Red Hat Gluster Storage algorithmically places and locates files. All of the necessary metadata is stored in extended attributes of files and directories.

All storage node servers in the cluster have the intelligence to locate any piece of data without searching in an index or querying another server. This provides fully parallel access to the data and ensures linear performance scaling. The performance, availability, and stability advantages of not using an external metadata server are significant and, in some cases, momentous.

GLOBAL NAMESPACE TECHNOLOGY

While many storage vendors contribute easier management of network storage to their namespace capability, the Red Hat Gluster Storage global namespace technology enables even greater capabilities.

STANDARDS-BASED FILE AND OBJECT STORE

With Red Hat Gluster Storage, all standard industry clients for file and object access are supported, including NFS, CIFS/SMB, and OpenStack Swift REST APIs. Applications accessing storage are not locked into any proprietary clients or closed interfaces, ensuring application portability.

CONCLUSION

Data archival processes can be time consuming and difficult depending on the source format of the data. Additionally, providing long-term retention is often times very costly as storage formats age and become obsolete. Red Hat Gluster Storage provides a number of key differentiators (above) that make it an ideal storage platform for data archival.



Red Hat RFI Response- Topic 3

Solicitation NNH15ZDA012L: Preparation for the Development of a Community-based Roadmap for NASA's Planetary Data Services

Date of Submission: January 4, 2016

National Aeronautics and Space Administration Point of Contact:

Dr. Michael New

michael.h.new@nasa.gov

Red Hat Point of Contact:

Eamon McCormick, Account Manager

Phone: [443.413.2719](tel:443.413.2719)

emccormi@redhat.com

Red Hat Inc. - Proprietary and Confidential

Treatment of Red Hat's Proposal

This response document contains Red Hat trade secrets, including but not limited to sensitive pricing information, and may not be disclosed to a third party without Red Hat's prior written consent. The exchange of information under this response shall be protected under any applicable agreement between the parties related to the treatment of confidential information. In the absence of such agreement, the recipient of this response shall protect the information contained herein in the same manner as it protects its own confidential information and may not be shared with any third party.

This is a proposal for a potential business relationship, and is neither a contract nor an offer that could be accepted by the other party. If the parties decide to enter into a business relationship, they will memorialize the terms of that relationship in a separate agreement between them or their respective business partners.

This document is Copyright 2015 by Red Hat, Inc. All rights reserved.

Contents

EXECUTIVE SUMMARY.....	1
<i>Red Hat Response</i>.....	2
Red Hat JBoss Fuse Functional Components.....	2
Red Hat JBoss Fuse Connectors.....	5

EXECUTIVE SUMMARY

The Planetary Data System identifies as key capabilities “flexibility for data submission, but also standardized structures for user to find and consume that data. In the current PDS, potential data providers connect with the appropriate “Discipline Nodes” and are then provided templates defining how data structures are to be submitted. The provider and node then iterate over sample submissions until the process and format are validated.

This approach is obviously successful but puts a great deal of burden on both the nodes and providers. Red Hat proposes implementing a flexible integration capability, composed of Enterprise Integration Patterns and Messaging, to deliver more flexibility and fault tolerance to the process. Adopting this approach would reduce the amount of manual work required by both nodes and providers, increase the speed of processing, improve accuracy, and allow for the addition of new providers and data formats quickly.

Red Hat JBoss Fuse and Red Hat JBoss A-MQ are powerful integration and messaging technologies, bundled as a single package, that have solved wide ranging integration and processing challenges. These technologies can handle incredibly large volumes of data, data in multiple formats, and many different transport protocols. Configured properly, they can intelligently route and transform the data delivered to, and ingested by, the PDS. Beyond the overall system improvements that can be delivered, adopting these technologies would allow for a reduction in custom code, and therefore custom code maintenance, in the PDS system.

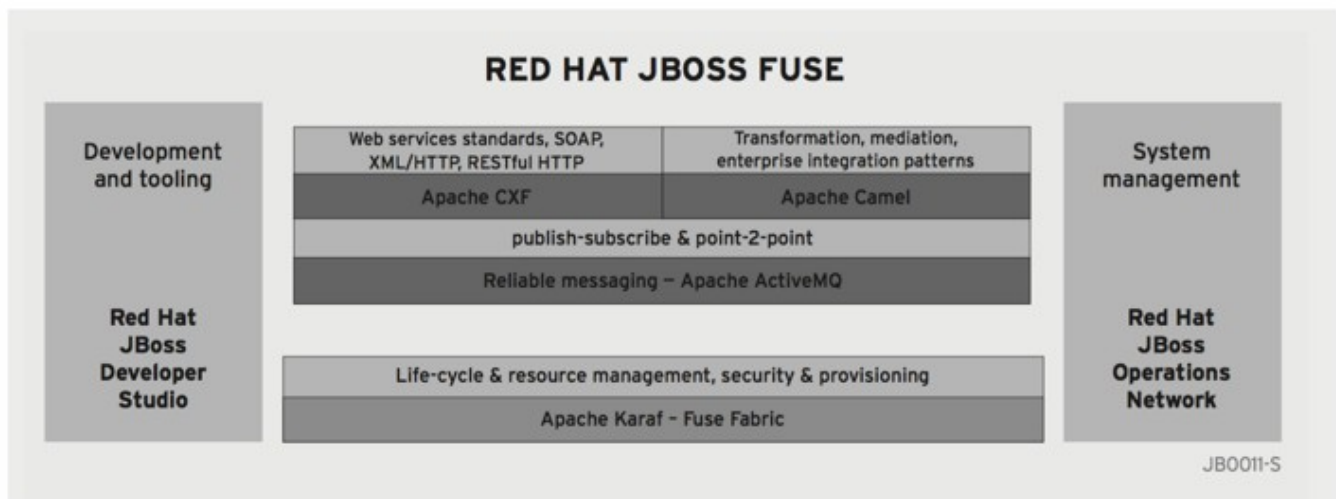
Red Hat Response

Enterprises are becoming increasingly connected, enabling digital transformations, increasing productivity, and facilitating rapid innovation. Red Hat JBoss Fuse, a lightweight integration platform, reduces the pain of connecting data sources, applications, services, processes, and devices for comprehensive and efficient solutions. JBoss Fuse includes the popular and versatile Apache Camel project, an implementation of the most commonly used enterprise integration patterns. With integration patterns and over 150 connectors ready to use, JBoss Fuse supports integration among internal systems, devices, and data sources, but also with those external to an organization, datacenter or set of datacenters.

Some integration challenges require comprehensive integration capabilities, while others need fast-to-develop, easy-to-manage integration platforms with small footprints. Some integration challenges require both. JBoss Fuse can be deployed and easily managed in any configuration, so it is possible to support multiple use cases. JBoss Fuse allows you to deploy a network of configurations across your infrastructure -- on premise, in the cloud, or in a hybrid environment -- to modernize your integration architecture and build a future-ready, connected solution.

JBoss Fuse also makes it possible to extend integrations with other capabilities such as real time business rules processing, business process management (BPM), distributed data caching, and more for a holistic connected solution. With this cost-effective, modular, lightweight, and cloud-ready integration platform, businesses can integrate in a faster, smarter way.

Red Hat JBoss Fuse Functional Components



The functional components of Red Hat JBoss Fuse include:

Container: The foundation of JBoss Fuse is a container. Manage large numbers of distributed containers with Fuse Fabric (based on Apache Karaf). Alternatively, deploy the JBoss Fuse based integration applications on Java EE-based Red Hat JBoss Enterprise Application Platform (JBoss EAP).

Integration framework: Use a standard method of notation and a high-level, domain-specific language to go from diagram to implementation with minimal coding. This layer is based on Apache Camel and includes over 150 connectors.

Web services framework: Turn any application or system into a service for inclusion in your service-based architecture. Service enablement technology is based on Apache CXF.

Reliable messaging: Red Hat JBoss A-MQ, a secure, standards-based message broker based on Apache ActiveMQ, easily extends your datacenter to the Internet of Things.

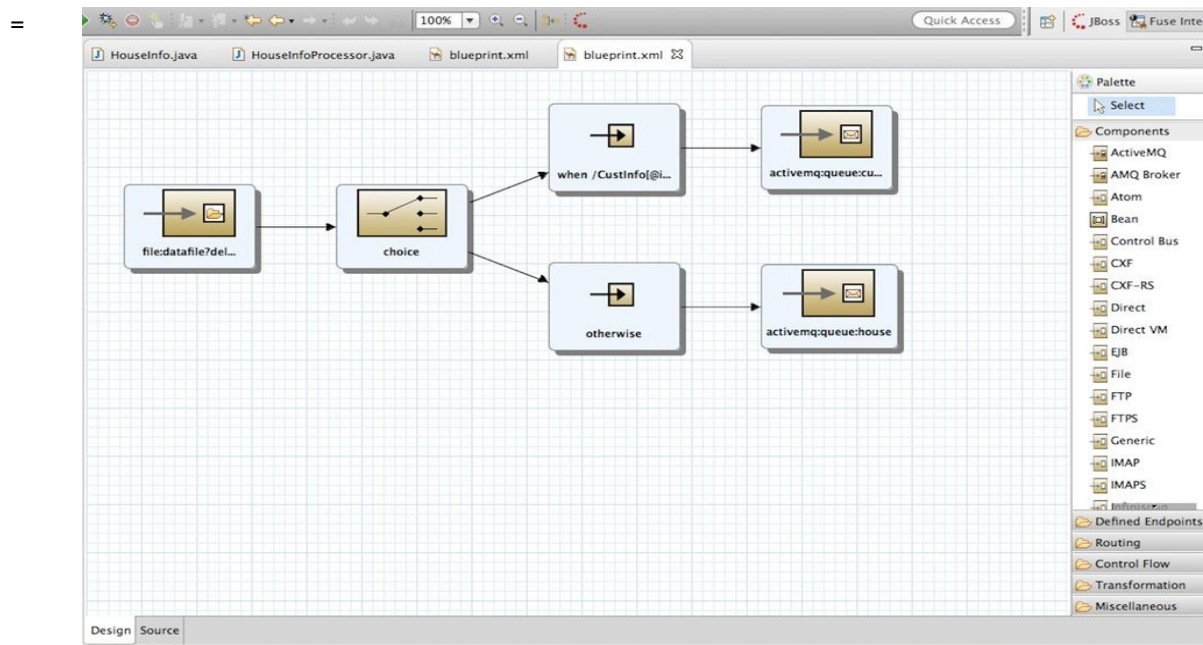
Development and tooling: Red Hat JBoss Developer Studio, with Fuse IDE, supports JBoss Fuse with intuitive tooling to help you with development. Drag and drop prebuilt integration patterns, add transformations and connectors, and visually map data to quickly create integration services. Debug integration services from the same tool for better quality.

API foundation: Create APIs that encapsulate the complexity of integrating and connecting multiple applications. Share APIs for easier collaboration with your suppliers, customers, and partners.

Management and monitoring: Manage your JBoss Middleware infrastructure with Fabric Management Console. Monitor your infrastructure with Red Hat JBoss Operations Network.

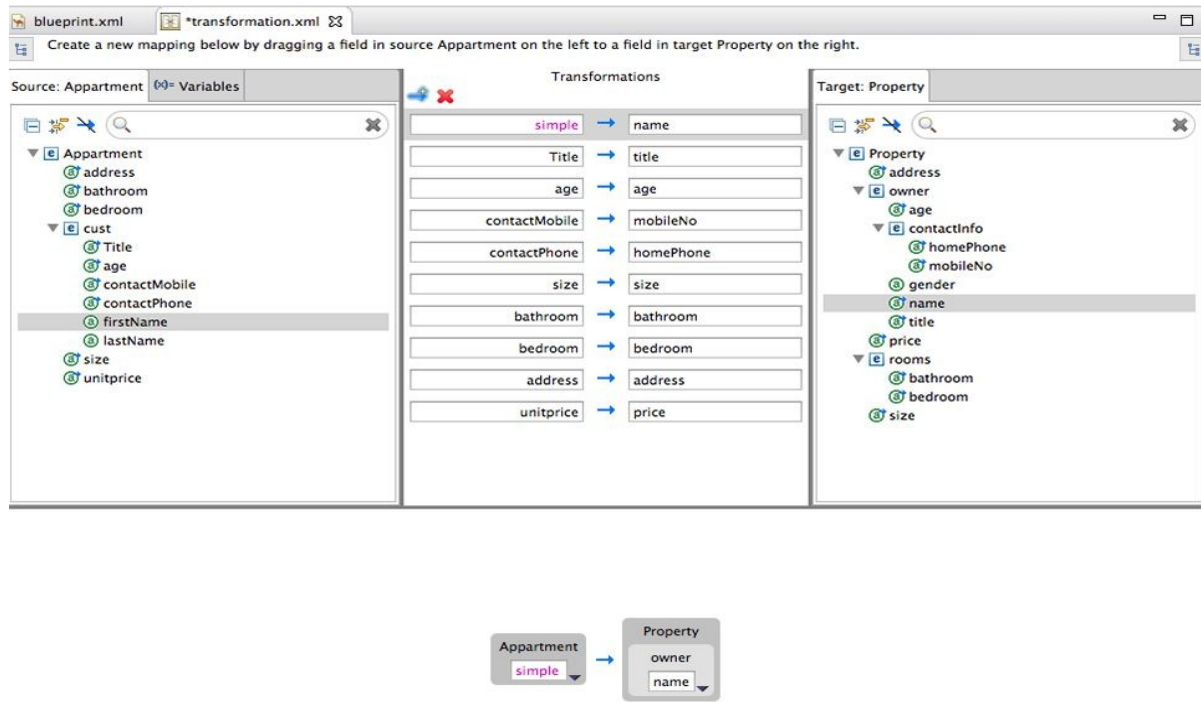
Red Hat JBoss Fuse includes the same integration capabilities (Apache Camel, Apache ActiveMQ and Apache CXF) found in Apache ServiceMix and expands those capabilities with Fuse Fabric and JBoss Operations Network for simplified management and monitoring of different deployment architectures.

As shown below, creating a new integration pattern is as simple as modeling that pattern, and specifying properties around each point in the integration. The visual example below shows ingesting of a data file (format defined by the integration owner), a decision based on file content, and routing of information to a specific message broker based on that content. From the broker, that file or the information collected from the file could be sent to a relational database, file system, Hadoop-based repository, or another big data repository.



One of the challenges that come with system and data integration is that the component systems often work with different data formats. You cannot simply send messages from one system to another without translating it into a format (or language) recognized by the receiving system. Data transformation is the term given to this translation.

The Fuse Transformation tooling provided with JBoss Fuse is a GUI to assist developers in implementing data translations as part of Camel routes. Data can be transformed from a large number of source formats to an equally large number of target formats, thus allowing for more flexibility at the integration layer. This flexibility reduces the burden on both providers and consumers of data.



JBoss Fuse supports an array of:

Web services standards and APIs: JAX-RS; JAX-WS; JSR 181; SAAJ; SOAP; MTOM; WSDL; WS*

Messaging standards and APIs: JMS; AMQP

Transport Protocols: HTTP/S; TCP/IP over SSL; MQTT; OpenWire; and STOMP





























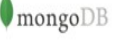



Additional standards and APIs: JDBC; Streaming API for XML Processing; JAF; JPA, and more:

<https://access.redhat.com/articles/375743?tour=7>

Red Hat JBoss Fuse Connectors

Additionally, JBoss Fuse delivers a set of “connectors” that provide easy integration of your systems, devices, and applications with third party APIs. Please view a sample below. The complete list is available here:

<http://www.jboss.org/products/fuse/connectors/>.

 <p>File Access file systems to easily move transformed data to files or read data from files for further processing</p>	 <p>JDBC Easily access or change application information stored in a database through JDBC</p>	 <p>Twitter Allow direct, polling, or event-driven consumption of timelines, users, trends, and direct messages using the Twitter API.</p>	 <p>SAP NetWeaver Easily consume SAP services through HTTP transport using the SAP NetWeaver Gateway</p>	 <p>AMQP Consume and publish messages with any AMQP 1.0 compliant messaging platform</p>	 <p>DropBox Read, search, write or manipulate files in DropBox using DropBox API's</p>	 <p>JBoss Data Grid Read or write to a fully-supported distributed cache and data grid for faster integration services</p>	 <p>ActiveMQ Easily consume and publish messages from JMS queues or topics</p>
 <p>HL7 Exchange, integrate, share, retrieve Electronic health information based on HL7 model using the HAPI library</p>	 <p>SAP Enable outbound and inbound communication to and from SAP systems using synchronous remote function calls, sRFC.</p>	 <p>Facebook Easily interface with Facebook API's to securely access or change social updates</p>	 <p>Salesforce Easily interface with Salesforce API's for inbound as well as outbound communication from Salesforce</p>	 <p>GHTTP Provide connectivity to the GAE URL fetch service and can also be used to receive messages from servlets</p>	 <p>GTask Supports asynchronous message processing on GAE using the task queueing service as a message queue.</p>	 <p>MQTT Consume IOT or M2M messages from MQTT compliant message brokers like Apache ActiveMQ or Mosquitto</p>	 <p>Google Drive Read, search, write or manipulate files on Google Drive using Google Drive API's</p>
 <p>GLogin Used by Camel applications outside Google App Engine (GAE) for programmatic login to GAE applications.</p>	 <p>Yammer Enable interaction with the Yammer enterprise social network.</p>	 <p>LinkedIn Easily retrieve LinkedIn user profiles, connections, groups, posts, and more using LinkedIn REST API's</p>	 <p>Amazon Web Services Easily interface with Amazon Web Services (AWS) like database, email, storage, messaging</p>	 <p>Box Read, write, manage, collaborate on files in Box using Box API's</p>	 <p>Olingo Expand your integration services to interface with Open Data Format (ODF) compliant data services</p>	 <p>LevelDB To provide persistent support for various Camel features such as Aggregator in Apache Camel Development Guide.</p>	 <p>CouchDB For enabling treatment of CouchDB instances as a producer or consumer of messages.</p>
 <p>GMail Supports sending of emails via the GAE mail service.</p>	 <p>FTP Easily read or write files to remote file systems using FTP or SFTP protocols.</p>	 <p>Geocoder Lookup geocodes (latitude and longitude) for a given address, or perform reverse look-up.</p>	 <p>GAuth Used by web applications for implementing a Google-specific OAuth consumer</p>	 <p>MongoDB Interface with NoSQL MongoDB database and collections</p>	 <p>Exec Extend solution by executing system commands to start other processes or jobs.</p>	 <p>Cometd Push exchanges or messages directly into the browser using an AJAX based mechanism.</p>	 <p>Class Easily bind message exchanges to java class or Plain old java objects (POJO)</p>



Red Hat RFI Response- Topic 7

Solicitation NNH15ZDA012L: Preparation for the Development of a Community-based Roadmap for NASA's Planetary Data Services

Date of Submission: January 4, 2016

National Aeronautics and Space Administration Point of Contact:

Dr. Michael New

michael.h.new@nasa.gov

Red Hat Point of Contact:

Jon Shallant, Account Manager

Phone: 703.629.0564

jshallant@redhat.com

Treatment of Red Hat's Proposal

This response document contains Red Hat trade secrets, including but not limited to sensitive pricing information, and may not be disclosed to a third party without Red Hat's prior written consent. The exchange of information under this response shall be protected under any applicable agreement between the parties related to the treatment of confidential information. In the absence of such agreement, the recipient of this response shall protect the information contained herein in the same manner as it protects its own confidential information and may not be shared with any third party.

This is a proposal for a potential business relationship, and is neither a contract nor an offer that could be accepted by the other party. If the parties decide to enter into a business relationship, they will memorialize the terms of that relationship in a separate agreement between them or their respective business partners.

This document is Copyright 2015 by Red Hat, Inc. All rights reserved.

Contents

EXECUTIVE SUMMARY.....	1
<i>Red Hat Response</i>.....	2
SOLUTION OVERVIEW.....	2
RED HAT GLUSTER STORAGE DIFFERENTIATORS.....	2
SOFTWARE-ONLY STORAGE.....	2
OPEN SOURCE SOFTWARE.....	2
SUPPORT FOR STORAGE CO-RESIDENT AND DATA LOCALITY AWARE APPLICATIONS.....	3
USER SPACE.....	3
MODULAR, STACKABLE ARCHITECTURE.....	3
DATA STORED IN NATIVE FORMATS.....	3
NO EXTERNAL METADATA SERVER.....	3
GLOBAL NAMESPACE TECHNOLOGY.....	3
STANDARDS-BASED FILE AND OBJECT STORE.....	4
CONCLUSION.....	4

EXECUTIVE SUMMARY

Red Hat is the world's leading provider of open source solutions, using a community-powered approach to provide reliable and high-performing cloud, virtualization, storage, operating system, and middleware technologies. Open source is a model for software developed by a community of diverse developers, resulting in higher quality, more secure, more easily integrated software at a vastly accelerated pace.

With a Red Hat subscription, there are no hefty up-front licensing fees, no support incident limits, no upgrade costs, and no hidden charges. Instead, NASA will receive unlimited support, software binaries, product documentation, updates, upgrades, bug fixes, and security patches all at one low, predictable price.

Red Hat has a proven track record managing data and enabling end users to get faster access to data, as shown in the following three case studies:

1. Cern is the European Organization for Nuclear Research, a conglomeration that spans 34 countries, 140 facilities and 100K machines. They needed to send information across their network but their homegrown system was brittle, prone to losing messages and could not scale. They used Red Hat JBoss Fuse to allow for fast, resilient system for sharing data across the different machines to improve research and collaboration.

2. The Federal Aviation Administration (FAA) needed a reliable, cost effective means for data exchange between themselves, industry and airline partners. The platform needed to be easy for partners to work with, and thus had to accommodate numerous protocols, standards and data formats to allow for flexible yet secure data exchange. The platform must also simultaneously handle the increased capacity and future demand. The FAA utilizes Red Hat to provide all these capabilities to their strategic System Wide Information Management (SWIM) program while reducing interdependence on existing systems, making them more agile in their data exchange capabilities.

3. At the University of Reading, the Department of Meteorology needed a highly reliable, available, and scalable storage file system to manage data for its scientific research projects in weather, climate, and earth observation. With Red Hat Storage Server, the department now has an enterprise-grade product—backed by world-class service and support—that saves departmental IT staff valuable research time they used to spend on maintenance and administration tasks.

Red Hat and NASA have had an outstanding relationship for many years. Red Hat has presence at every NASA Center, and within many scientific and research organizations. NASA leverages not only Red Hat Enterprise Linux, but also many solutions from our Cloud, Middleware and Storage portfolio. Red Hat understands NASA's passion for open source solutions as we continue to support the organization in increasing productivity, enhancing security/compliance requirements and lowering the costs of doing business. Red Hat believes we are in the best position to meet many of the items listed in the RFI. We look forward to NASA's response and the possibility of adding value to this specific organization's mission.

Red Hat Response

SOLUTION OVERVIEW

During the last decade, enterprises have seen enormous gains in scalability, flexibility, and affordability as they have migrated from proprietary, monolithic server architectures to virtualized, open source, standardized, and commoditized servers.

Unfortunately, the same has not been true in the storage industry. Proprietary, monolithic, and scale-up solutions, which dominate the storage industry today, do not deliver the outcomes of a modern, software-defined datacenter. They do not deliver scalability, flexibility, or the economics that datacenters, workloads, and cloud computing environments need in today's hyper-growth, virtualized, and increasingly cloud-based world. Red Hat Gluster Storage addresses this gap.

Red Hat Gluster Storage is an open, software-defined storage solution for private, public, and hybrid cloud environments. Based on the open source project GlusterFS, Red Hat Gluster Storage provides a distributed scale-out file systems technology to meet the needs of unstructured, semi-structured and big data storage environments.

Red Hat Gluster Storage enables organizations to combine large numbers of commodity storage and compute resources into a high-performance and centrally managed pool of storage. Capacity and performance can scale linearly and independently on-demand, from a few terabytes to petabytes and beyond, using both on-premise commodity hardware and the public cloud compute and storage infrastructure. As such, Red Hat Gluster Storage will improve search capabilities of the PDS, allowing researchers improved access to data products and metadata.

RED HAT GLUSTER STORAGE DIFFERENTIATORS

SOFTWARE-ONLY STORAGE

One of the cornerstone principles defining Red Hat Gluster Storage is that storage should be considered a software problem. Locking customers into one particular storage hardware vendor or one particular hardware configuration cannot solve today's storage problems. As a software-only storage solution, Red Hat Gluster Storage has been designed to work with a wide variety of industry-standard commodity storage, networking, and compute servers.

OPEN SOURCE SOFTWARE

Red Hat Gluster Storage delivers functionality by embracing the open source model. Based on the open source project GlusterFS, a distributed scale-out file system technology, Red Hat Gluster Storage takes many "upstream" projects in the community and packages it with Red Hat Enterprise Linux, for a stable and enterprise-grade storage offering. As a result, Red Hat Gluster Storage users benefit from a worldwide community of developers. These developers are constantly testing the product in a wide range of environments and workloads and providing continuous feedback.

SUPPORT FOR STORAGE CO-RESIDENT AND DATA LOCALITY AWARE APPLICATIONS

Red Hat Gluster Storage is ideal for running applications that need to run co-resident with storage, like Apache Hadoop map-reduce. These types of applications must avoid the extra network hop between the compute node and the storage server if possible to improve response times for analytics workloads.

The software-defined characteristic of Red Hat Gluster Storage makes it amenable to these workload types.

USER SPACE

Unlike traditional file systems, Red Hat Gluster Storage operates in user space. This significantly eases the process of installing and upgrading the product. It also enables expedited delivery of features that do not depend on kernel merge windows. And it means that users who contribute to GlusterFS do not need specialized kernel expertise.

MODULAR, STACKABLE ARCHITECTURE

Red Hat Gluster Storage is designed using a modular and stackable architecture approach. Configuring Red Hat Gluster Storage for highly specialized environments is a simple matter of including or excluding particular modules.

DATA STORED IN NATIVE FORMATS

With Red Hat Gluster Storage, data is stored on disks using native formats (e.g., XFS). The product has implemented various self-healing processes for data and therefore is extremely resilient. Furthermore, files are naturally readable without Red Hat Gluster Storage. Therefore, if a customer migrates away from Red Hat Gluster Storage, his/her data is still completely usable without any required modifications or data migration.

NO EXTERNAL METADATA SERVER

In a scale-out system, one of the biggest challenges is keeping track of the logical and physical data locations. Most distributed systems solve this problem by creating a separate index with file names and location metadata that usually resides in a server process, and is often referred to as a metadata server.

Unfortunately, this results in both a central point of failure and a huge performance bottleneck. Red Hat Gluster Storage does not create, store, or use a separate index of metadata that needs to be externally stored. Instead, Red Hat Gluster Storage algorithmically places and locates files. All of the necessary metadata is stored in extended attributes of files and directories.

All storage node servers in the cluster have the intelligence to locate any piece of data without searching in an index or querying another server. This provides fully parallel access to the data and ensures linear performance scaling. The performance, availability, and stability advantages of not using an external metadata server are significant and, in some cases, momentous.

GLOBAL NAMESPACE TECHNOLOGY

While many storage vendors contribute easier management of network storage to their namespace capability, the Red Hat Gluster Storage global namespace technology enables even greater capabilities.

STANDARDS-BASED FILE AND OBJECT STORE

With Red Hat Gluster Storage, all standard industry clients for file and object access are supported, including NFS, CIFS/SMB, and OpenStack Swift REST APIs. Applications accessing storage are not locked into any proprietary clients or closed interfaces, ensuring application portability.

CONCLUSION

Data archival processes can be time consuming and difficult depending on the source format of the data. Additionally, providing long-term retention is often times very costly as storage formats age and become obsolete. Red Hat Gluster Storage provides a number of key differentiators (above) that make it an ideal storage platform for data archival.

